

Claims

WHAT IS CLAIMED IS:

1. A method for multi-level distributed speech recognition
5 comprising:

providing an audio command to a first speech recognition
engine and at least one second speech recognition engine;
recognizing the audio command within the first speech
recognition engine to generate at least one first recognized
10 audio command, wherein the at least one first recognized
audio command has a corresponding first confidence
value; and
recognizing the audio command within the at least one second
speech recognition engine, independent of recognizing the
15 audio command by the first speech recognition engine, to
generate at least one second recognized audio command,
wherein the at least one second recognized audio
command has a corresponding second confidence value.

20 2. The method of claim 1 further comprising:

selecting at least one recognized audio command having a
recognized audio command confidence value from the at
least one first recognized audio command and the at least
one second recognized audio command based on the at

least one first confidence value and the at least one second confidence value.

3. The method of claim 2 further comprising:
5 prior to selecting at least one recognized audio command, weighting the at least one first confidence value by a first weight factor and weighting the at least one second confidence values by a second weight factor.

10 4. The method of claim 2 further comprising:
executing at least one operation based on the at least one recognized audio command.

15 5. The method of claim 2 further comprising:
verifying the at least one recognized audio command.

6. The method of claim 1 further comprising:
generating an error notification when the at least one first confidence value and the at least one second confidence values are below a minimum confidence level.
20

7. A method for multi-level distributed speech recognition comprising:

providing an audio command to a terminal speech recognition engine and at least one network speech recognition engine;

5

recognizing the audio command within the terminal speech recognition engine to generate at least one terminal recognized audio command, wherein the at least one terminal recognized audio command has a corresponding terminal confidence value;

10

recognizing the audio command within the at least one network speech recognition engine to generate at least one network recognized audio command, wherein the at least one network recognized audio command has a corresponding network confidence value; and

15

selecting at least one recognized audio command having a recognized audio command confidence value from the at least one terminal recognized audio command and the at least one network recognized audio command.

20

8. The method of claim 7 further comprising:

generating an error notification when the at least one terminal confidence value and the at least one network confidence value are below a minimum confidence level.

25

9. The method of claim 7 further comprising:

5 prior to selecting the at least one recognized audio command, weighting the at least one terminal confidence value by a terminal weight factor and the at least one network confidence value by a network weight factor.

10. The method of claim 7 further comprising:

10 filtering the at least one recognized audio command based on the at least one recognized audio command confidence value; and executing an operation based on the recognized audio command having the highest recognized audio command confidence value.

15 11. The method of claim 7 further comprising:

verifying the at least one recognized audio command to generate a verified recognized audio command; and executing an operation based on the verified recognized audio command.

20

12. An apparatus for multi-level distributed speech recognition comprising:

5 a first speech recognition means, operably coupled to an audio subsystem, for receiving an audio command and generating at least one first recognized audio command, wherein the at least one first recognized audio command has a first confidence value;

10 a second speech recognition means, operably coupled to the audio subsystem, for receiving the audio command and generating, independent of the first speech recognition means, at least one second recognized audio command, wherein each of the at least one second recognized audio command has a second confidence value; and

15 a means, operably coupled to the first speech recognition means and the second speech recognition means, for receiving the at least one first recognized audio command and the at least one second recognized audio command.

13. The apparatus of claim 12 further comprising:

20 a dialog manager operably coupled to the means for receiving, wherein the means for receiving selects at least one recognized audio command having a recognized confidence value from the at least one first recognized audio command and the at least one second recognized audio command based on the at least one first confidence

25

value and the at least one second confidence value, wherein the selected at least one recognized audio command is provided to the dialog manager.

5 14. The apparatus of claim 12 wherein the dialog manager determines a dialog manager audio command from the at least one recognized audio command based on the at least one recognized audio command confidence levels and wherein the dialog manager executes an operation in response to the dialog manager audio command.

10 15. The apparatus of claim 14 further comprising: wherein the dialog manager accesses a content server and retrieves encoded information in response to the dialog manager audio command.

15 16. The apparatus of claim 15 further comprising: a speech synthesis engine operably coupled to the dialog manager, wherein the speech synthesis engine receives speech encoded information from the dialog manager and generates speech formatted information.

20 17. The apparatus of claim 16 wherein the audio subsystem is operably coupled to the speech synthesis engine, wherein the audio subsystem receives the speech formatted information and provides an output message.

18. The apparatus of claim 17 wherein when the comparator provides the dialog manager with an error notification, the output message is an error statement.

5

19. A system for multi-level distributed speech recognition comprising:

a terminal speech recognition engine operably coupled to a microphone and coupled to receive an audio command

5 and generate at least one terminal recognized audio command, wherein the at least one terminal recognized audio command has a corresponding terminal confidence value;

at least one network speech recognition engine operably coupled

10 to the microphone and coupled to receive the audio command and generate at least one network recognized audio command, independent of the terminal speech recognition engine, wherein the at least one network recognized audio command has a corresponding network confidence value;

15 a comparator operably coupled to the terminal speech recognition engine operably coupled to receive the at least one terminal recognized audio command and further operably coupled to the at least one network speech recognition engine operably coupled to receive the at least one network recognized audio command; and

20 a dialog manager operably coupled to the comparator, wherein the comparator selects at least one recognized audio command having a recognized confidence value from the at least one terminal recognized audio command and the

25

at least one network recognized audio command based on
the at least one terminal confidence value and the at least
one network confidence value, wherein the selected at
least one recognized audio command is provided to the
5 dialog manager.

20. The system of claim 19 wherein the dialog manager
determines a dialog manager audio command from the at least one
recognized audio commands based on the at least one recognized
10 audio command confidence levels and wherein the dialog manager
executes an operation in response to the dialog manager audio
command.

21. The system of claim 20 further comprising:
15 wherein the dialog manager accesses a content server and
retrieves encoded information in response to the dialog
manager audio command.

22. The system of claim 21 further comprising:
20 a speech synthesis engine operably coupled to the dialog
manager, wherein the speech synthesis engine receives
speech encoded information from the dialog manager and
generates speech formatted information; and
a speaker operably coupled to the speech synthesis engine,
25 wherein the speaker receives the speech formatted

information and provides an output message.